



INFORMATION EXTRACTION TOOL (IET)

**John Baumler (baumlerj@rl.
AFRL/IFEB
315-330-3979 (DSN587-397
27 Nov 2000**



Background

- DODIIS AMHS provides message archiving, retrospective search and message generation/review capabilities for organizational messages (DMS Will Lack This Capability)
- 497IOG is the DODIIS Executive Agent for DODIIS AMHS
 - 26 legacy DODIIS AMHS sites
 - 2100/4100 hardware requires recap over next two years per COMPAQ
 - 5 DAWS Message Front End (DMFE) sites maintained by AFRL

Background (continued)



- Dec 98 - 497IG/IND requested AFRL develop a low cost automated message handling capability for the SCI DMS-era architecture
 - Add BROADSWORD as user interface
 - Provide migration path to DMS user agent (Outlook)
 - Support DODIIS Target Operating Systems (Solaris/Windows NT)
 - Selection Based on DMS Implementation, Functional Equivalence and Life Cycle Cost Analysis
 - Functional Requirements Comparison Available on IET Web Page



Information Extraction Tool (IET)

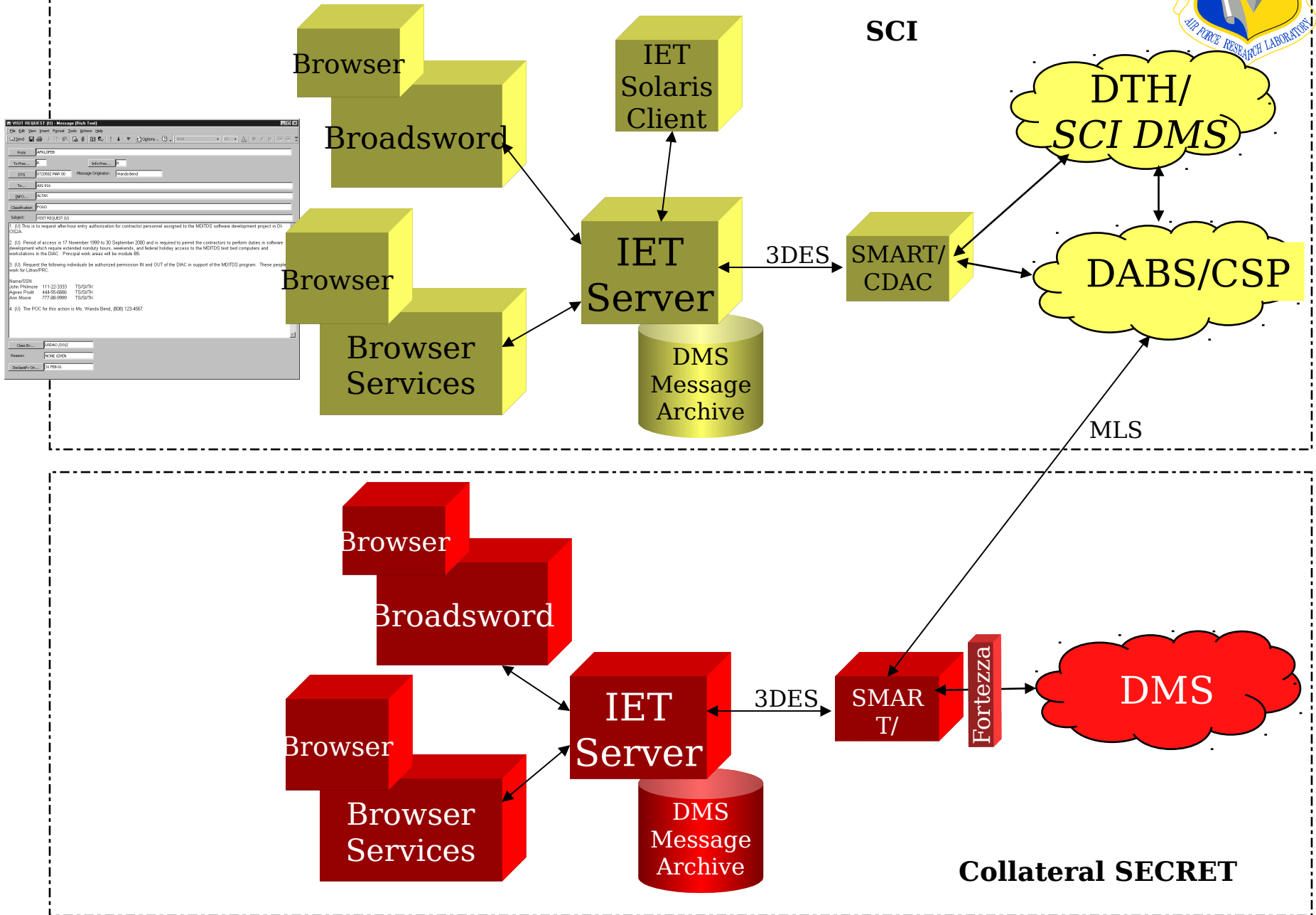
- Value added profiling and search capabilities for the intelligence analyst; supplements DMS architecture
- Provides analyst with access to archived messages (and enclosures-v1.1)
- Provides transitional message generation functionality
- Integrates into DMS-era architecture by interfacing with CSP/CDAC
- Maximizes use of existing accredited interfaces
- Supports local and remote user base
- Serves as platform for future integration of advanced information extraction technologies



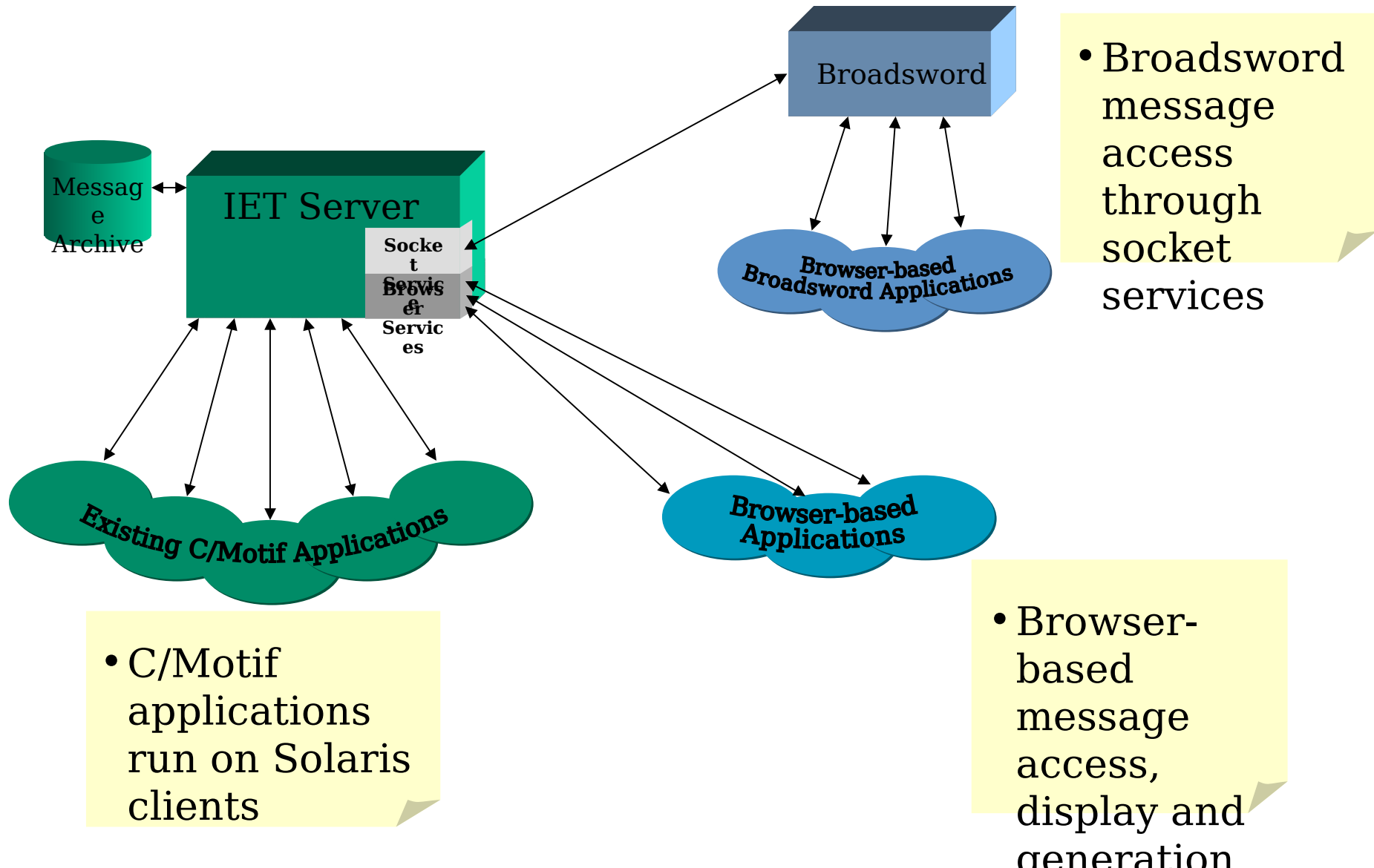
Information Extraction Tool (IET)

- DMFE provides IET core services
 - Sun Solaris server environment, Sybase DBMS
- Adds Browser-based interface
- BROADSWORD plug-in available for profiling/retrospective search

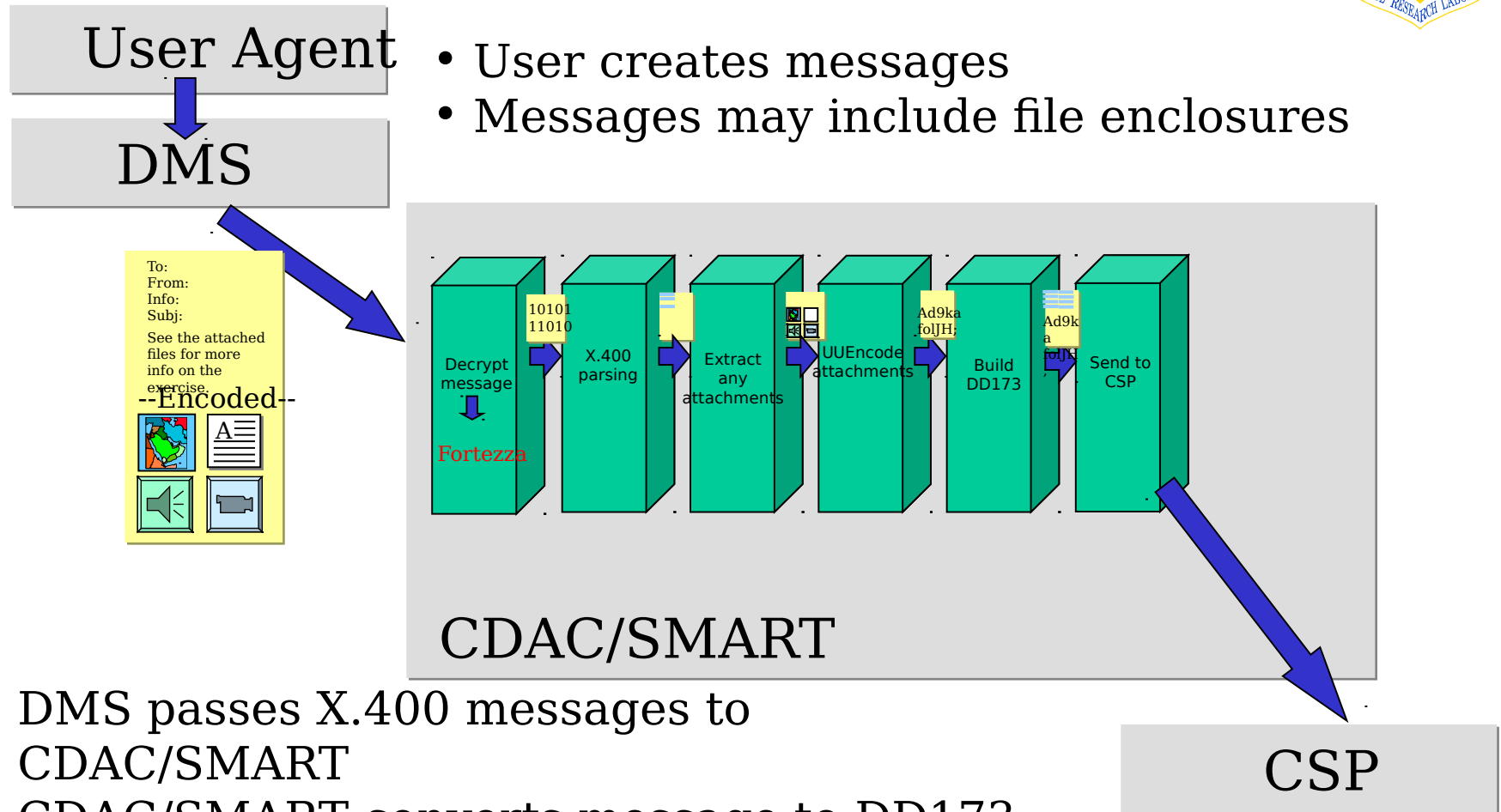
IET as Part of DMS Architecture



IET Client Environment



Message Flow - CDAC/SMART



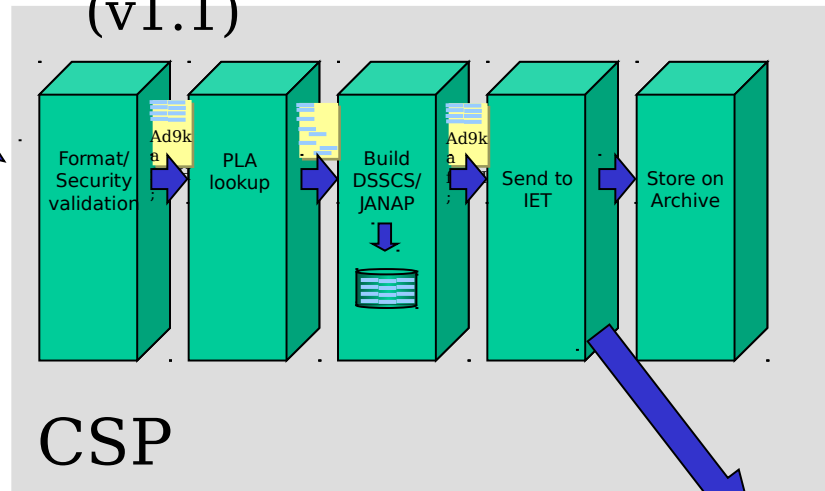
- DMS passes X.400 messages to CDAC/SMART
- CDAC/SMART converts message to DD173 format
- Messages may include file enclosures
- Enclosures are encoded, attached to the message and passed to CSP

Message Flow - CSP/IET

CDAC/SMART

Ad9k
a
folJH
;

- CSP passes messages to IET
- Messages may include file enclosures (v1.1)



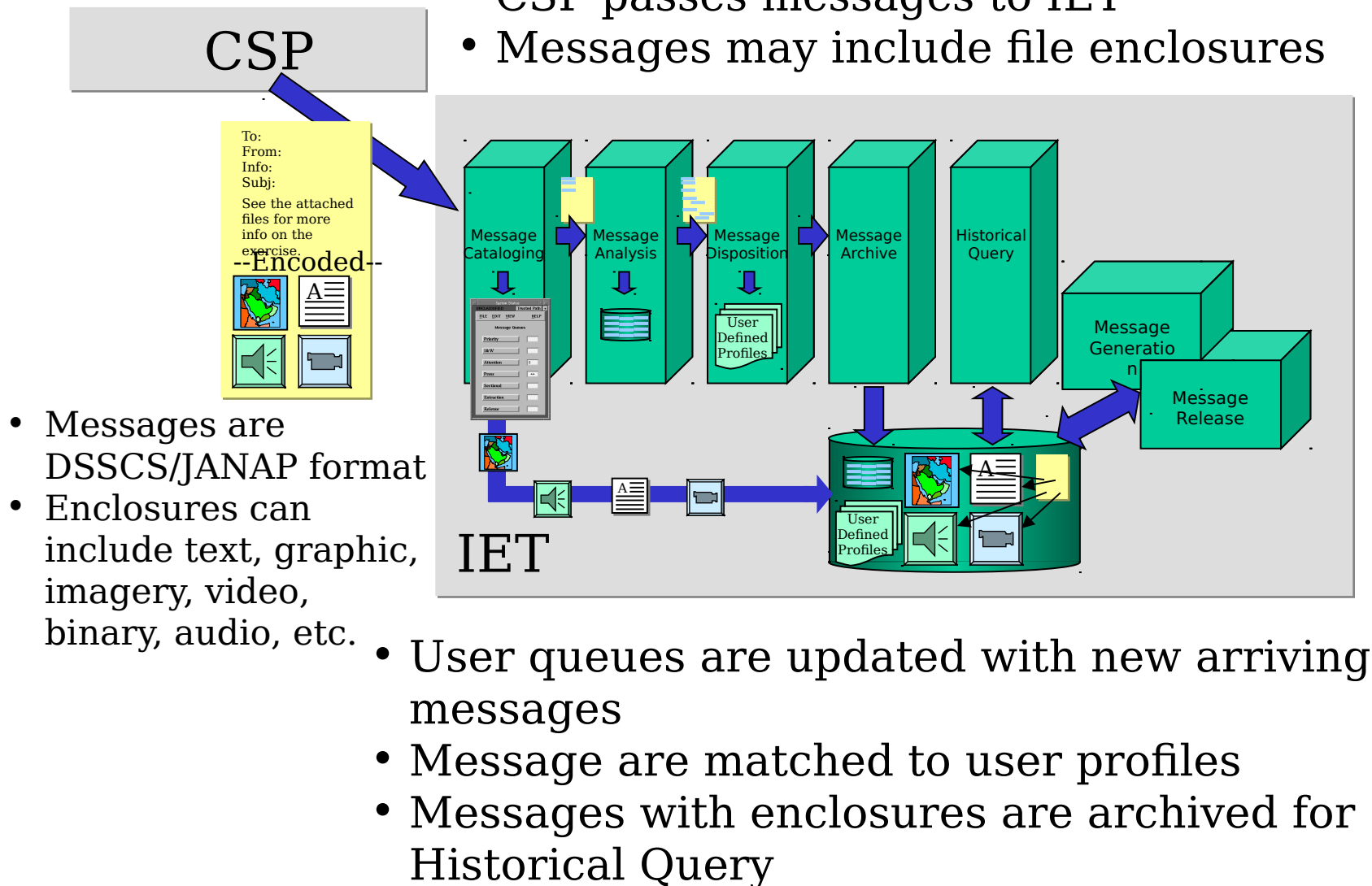
CSP

IET

- CSP converts DD173 messages to DSSCS/JANAP format
- Attachments are embedded in the message

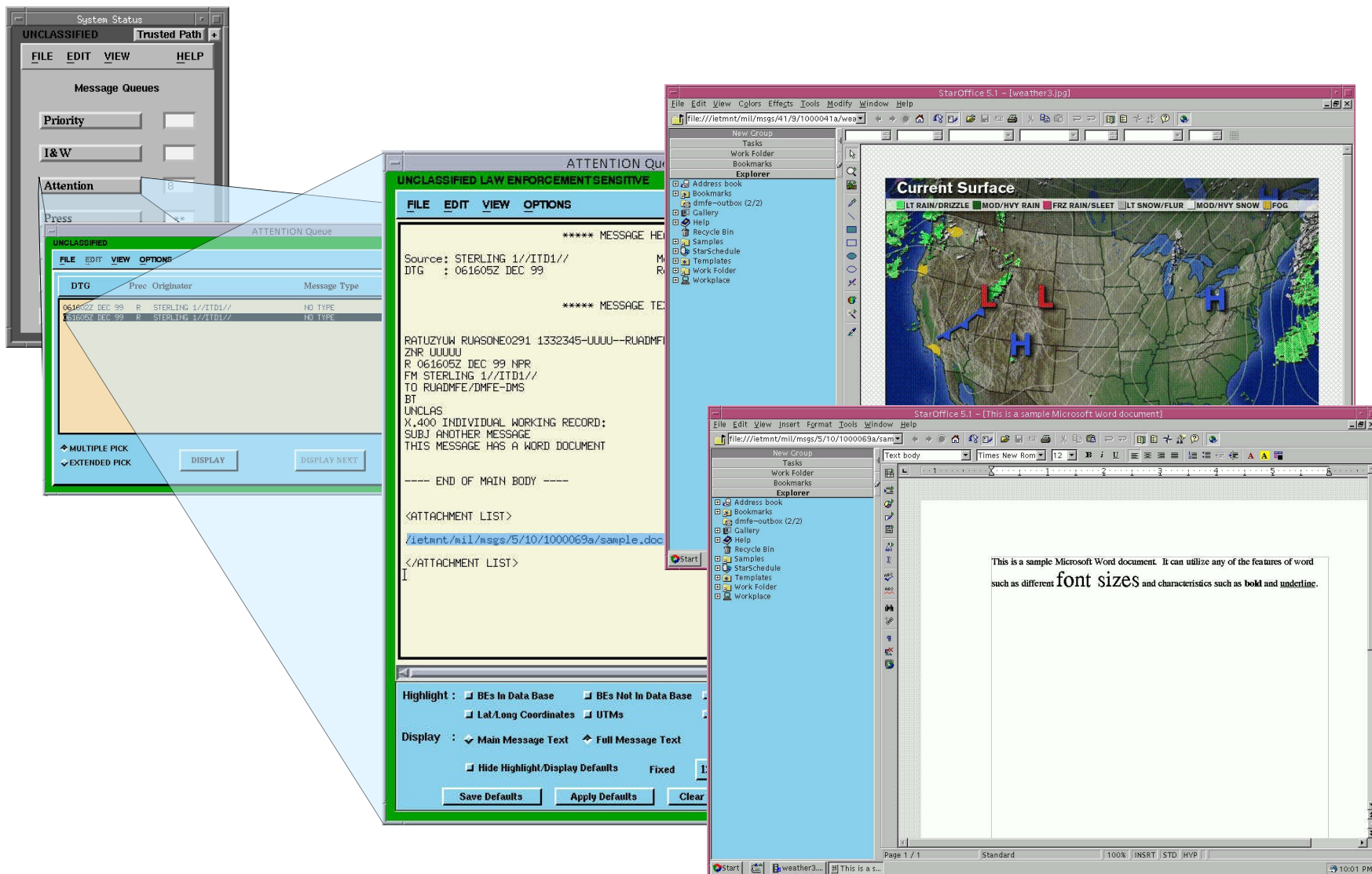
IET Processing

- CSP passes messages to IET
- Messages may include file enclosures





IET User Interface





IET Evolution

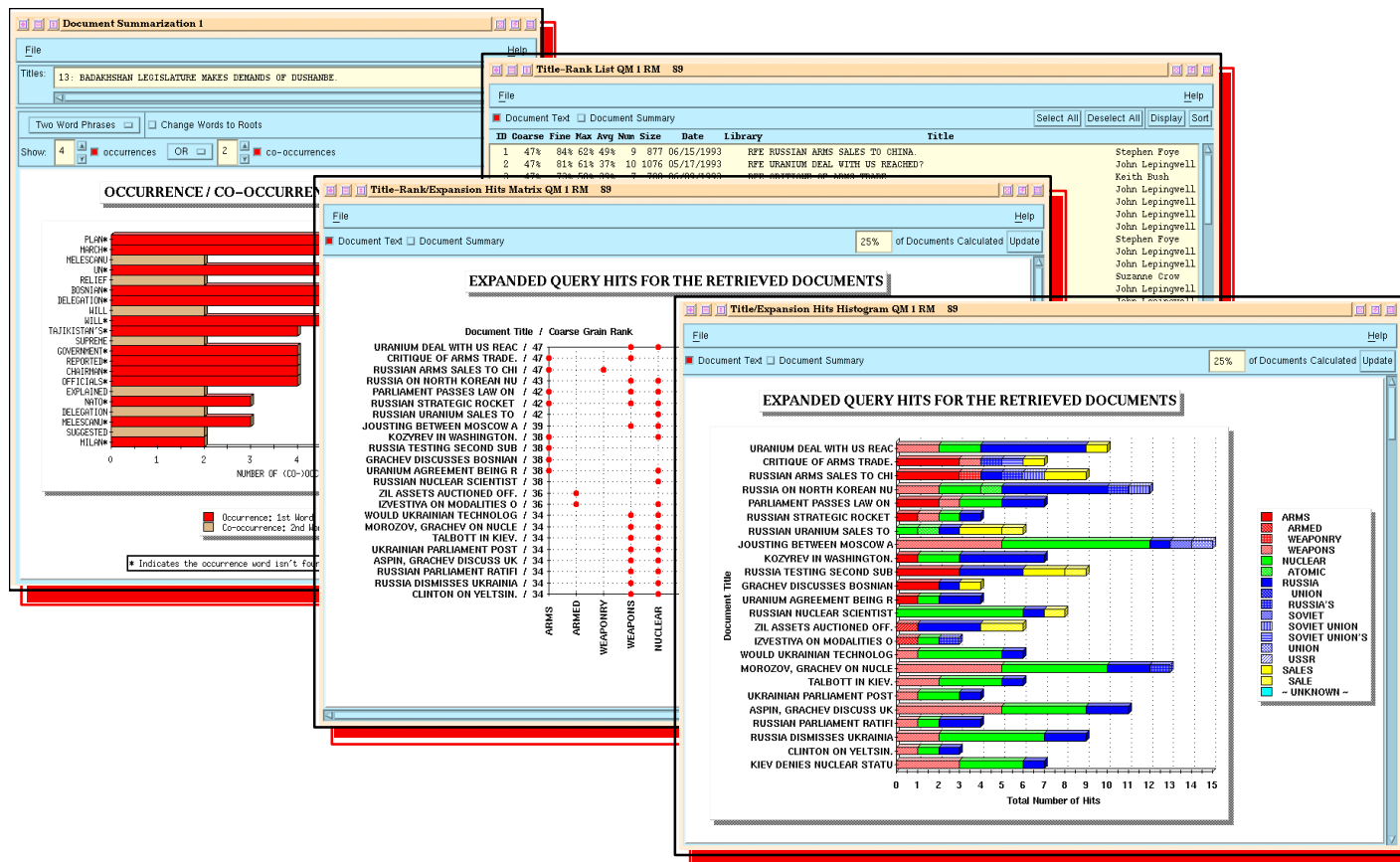
- IET v1.0 (Oct 2000 release)
 - Adds Browser interface for profiling, retrospective search and message generation/review
 - Adds BROADSWORD interface for profiling, retrospective search
 - Supports organizational messages
 - Functionally Equivalent to DODIIS AMHS
- IET v1.1 (in development-May 2001 DODIIS testing)
 - Provides capability to profile, store, perform retrospective search of messages with enclosures
 - DII COE Level 5 Compliance/AFDI integration
 - XML support
- IET v1.2 (late FY01)
 - Incorporate advanced message analysis tools (EMCARS)

Technology Insertion - Additional Advanced Features (v1.2)



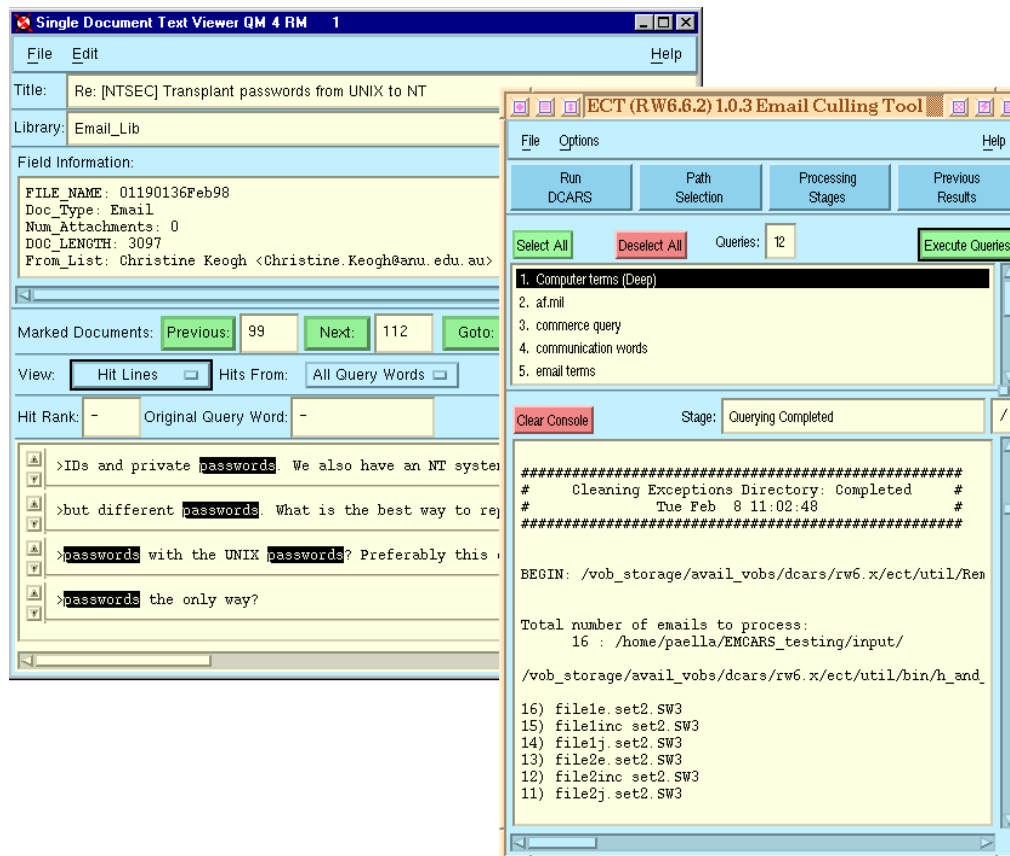
- Automatically handles attachments in over 200 of the most popular commercial proprietary formats
- Visualizations that aid in reducing the time it takes to find the documents that are most relevant to the analyst
- Sophisticated rapid-scan browsing and gisting capabilities for retrieved emails
- Built-in capability to categorize, save, load, edit and re-use both retrospective queries and profiles
- Options to construct and modify existing dictionaries/thesauri into new custom dictionaries/thesauri to enhance recall
- NASA Technical Thesaurus available to facilitate query formulation and provide high precision search (other dictionaries and thesauri are available)

Search Results Visualizations



Visualizations aid in reducing the time it takes to find the documents that are most relevant to the user

Sophisticated Rapid-Scan Browsing and Gisting



Sophisticated rapid-scan browsing and gisting capabilities for retrieved emails enhances rapid drill down

Deployment Options/Costs



- New standalone server (parts list available)
 - Small (\$20k) - 513MIB (Fort Gordon)
 - Medium (\$31k) - TRANSCOM, EUCOM, SPACECOM, 480IG
 - Large (\$67k) - SOUTHCOM, SOCOM, CENTCOM, JAC
- Reuse of existing Sun hardware (per discussions with individual sites)
- Installation cost approximately \$28k/per site
 - Includes software installation/configuration, data migration from legacy system (AMHS/DMFE), security accreditation testing, training
- No annual user fee for IET software maintenance/upgrades per 497IOG/IND



IET Activities

- IET v1.0 Beta II Testing
 - Successfully completed JITF testing, Aug 2000
 - Successfully completed Beta I testing at USSTRATCOM 5-7 Sep 2000
 - Conducting site demonstrations per 497IOG direction
- 497IOG Site Install, Oct 2000
- IET v1.1 Requirements Review, 497IOG, Jan 2001
- AFRL developing roadmap for information extraction technology insertion
 - Email Content and Retrieval System (EMCARS)
 - Generic Intelligence Processor (GIP)



Summary

- Information Extraction Tool provides DMS-era message handling functionality
 - Archiving
 - Profiling
 - Retrospective search
 - Message generation
 - Accommodates DMS messages with enclosures
- IET v1.0 (DMFE v7.0) ready for deployment Oct 2000
- 497IOG has recommended IET as replacement for DODIIS AMHS



AFRL/IFEB POCs

- Chief, IDHS Program Office - John Vona, 315-330-3601 (DSN587), vonaj@rl.af.mil
- Program Manager - John Baumler, 315-330-3979 (DSN587), baumlerj@rl.af.mil
- Lead Engineer - Mike Corelli, 315-330-7525 (DSN587), corellim@rl.af.mil
- AFRL/IFEB home page
 - Internet - <http://www.if.afrl.af.mil/programs/iet/>